

# PATENT APPLICATION

## I/O METHOD AND APPARATUS FOR OPTICAL STORAGE MEDIA

### Inventors:

Pontus Söderström  
Karl XI gatan 8a  
222 20 Lund  
046-211 56 35  
A Citizen of Sweden

Per Sten  
Linnég. 9  
224 60 Lund  
046-127479  
A Citizen of Sweden

Thu Le CHI  
Jägaregatan 41  
226 52 LUND  
046-151124  
A Citizen of Sweden

### Assignee:

**AXIS AB**  
Scheelevägen 34  
SE-223 63 Lund, Sweden  
A Swedish Corporation

CARY & KELLY, LLP  
1875 Charleston Road  
Mountain View, CA 94043  
Telephone (650) 316-4009

003021 10/EE/60

## **I/O METHOD AND APPARATUS FOR OPTICAL STORAGE MEDIA**

### **NOTICE REGARDING COPYRIGHTED MATERIAL**

5 A portion of the disclosure of this patent document contains material which is  
subject to copyright protection. The copyright owner has no objection to the facsimile  
reproduction by any one of the patent document or the patent disclosure as it appears  
in the public Patent Trademark Office file or records, but otherwise reserves all  
copyright rights whatsoever.

### **CROSS-REFERENCE TO RELATED APPLICATION**

10 This application claims priority from Provisional Application Numbers:  
60/169,745, entitled "PACKET CACHE" filed on 12/08/99 (Attorney Docket #  
AXISP001P); 60/170,078, entitled "FILE EXTENT CACHE" filed on 12/09/99  
(Attorney Docket # AXISP002P) which are incorporated herein by reference in their  
entirety.

### **BACKGROUND OF THE INVENTION**

#### **1. Field of the Invention**

This invention generally relates to high capacity storage devices more  
particularly the present invention relates to network attached high capacity optical  
storage devices.

#### **2. Description of the Related Art**

25 Optical media are used in a wide range of applications for storage of data. In  
the computer industry compact discs are used to store computer software. As optical  
storage densities increase compact disks are becoming more widely used in the  
entertainment, music and video industries. Compact disks are also used for archiving  
data and for storing reference documents.

30 An advantage of CD drives is that the compact disc (CD) has a high data  
storage capacity, e.g. 650 Meg and is portable. A CD may be individually removed  
from the drive replaced with another compact disc. Thus the user may have a  
collection of different discs which may be individually mounted on the CD drive as  
desired.

One disadvantage of CD-ROM drive technology is that the access times and transfer rates associated with CD-ROM drives are in general substantially slower than other comparable data storage technologies. An ordinary double speed CD drive commonly found in many of today's computers has an access time of 300 milliseconds. In contrast a comparable hard disk commonly found in many of today's computers has an access time of 10 milliseconds. Furthermore ordinary dynamic read-only memory (DRAM) found in many of today's computers has an access time of 60 ns. The amount of time required to read and write data to the CD drive is much greater than a corresponding time required to access other data storage technologies. With a large speed discrepancies between optical storage medium and other types of data storage there is a need to improve the performance, scalability, and availability of CD-ROM drives.

## SUMMARY OF THE INVENTION

The present invention provides a network enabled multi-user configurable optical storage server. The device is easy to install and does not require client or server software. The server can be configured across a network through a series of HTML based web pages, or graphical user interfaces (GUI) provided by the server.

The server can be accessed across a network by one or more users, terminals, or other network attached devices for substantially concurrent read/write access to one or more optical storage media. The server supports a comprehensive set of networking protocols and operating systems simultaneously. The server may support from one to hundreds of CD/DVD-ROMS.

In an embodiment of the invention an apparatus for providing I/O access to at least one data storage device across a network is disclosed. The apparatus includes a network module, a first and second stage and a logic. The network module is configured to couple to the network for sending and receiving data packets. The first and second stages serially couple the network module to the storage device. The logic is responsive to a received data packet from the network to serially move data from the received data packet from said network module through each of said first and said second stages to the at least one storage device. The logic is further responsive to a read request from the network to send a data packet to the network via said network module from whichever of said at least first and second stages and the at least one storage device includes the data.

In an alternate embodiment of the invention an apparatus for providing I/O access to at least one data storage medium across a network is disclosed. The apparatus includes a network module, at least one data storage device a hard drive and logic. The network module is configured to couple to the network for sending and receiving data packets. The at least one data storage device is coupled to the network module and the at least one data storage device for providing an input and an output of datum stored on the at least one data storage medium. The hard drive is coupled to the network module for the caching of at least one of file structures for the stored datum and file structures together with the corresponding datum stored on the at least one data storage medium. The logic is for determining a selected cache policy for the at least one data storage device based on a user selection and for caching on said hard drive a corresponding selected one of the file structures and file structures together with the corresponding datum responsive to the user input.

In an embodiment of the invention a method for providing I/O access to at least one data storage device across a network is disclosed. The method comprises the acts of:

coupling to the network for input and output of data;

coalescing data received from the network which corresponds with a selected file or data stream;

09733707 120800

storing said data aggregated in said act of aggregating.

coupling to the network for sending and receiving data packets;

10            caching at least one of file structures for the stored datum and file structures  
together with the corresponding datum stored on the at least one data storage medium  
responsive to a user selection.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, wherein like reference numerals  
5 designate like structural elements, and in which:

FIG. 1 shows a plurality of devices coupled across the network with an optical storage server.

FIG. 2 is a detailed hardware block diagram of the server shown in FIG. 1 including graphical user interfaces stored therein for server setup over the network.

10 FIGS. 3A-F are combined hardware block diagrams of the server and optical storage media shown in FIG. 1 during a write phase of I/O operation.

FIGS. 4A-B are process flow diagrams corresponding with the writing and reading of data to an optical or DVD media as shown in FIGS. 3A-F.

15 FIG. 5 is a hardware block diagram of the server, optical and electronic storage mediums shown in FIG. 1.

FIG. 6A-C show the data structures associated with both the CD and hard drives shown in FIG. 5.

20 FIGS. 7AB are process flow diagrams corresponding with the boot phase and operational phase of operation of the server shown in FIGS. 5-6AC.

## **DETAILED DESCRIPTION OF THE INVENTION**

The present invention provides a network enabled multi-user configurable optical storage server. The device is easy to install and does not require client or server software. The server can be configured across a network through a series of  
25 HTML based web pages, or graphical user interfaces (GUI) provided by the server. The server can be accessed across a network by one or more users, terminals, or other network attached devices for substantially concurrent read/write access to one or more optical storage media. The server supports a comprehensive set of networking protocols and operating systems simultaneously. The server may support from one to

09733707-120800

hundreds of CD/DVD-ROMS. In an embodiment of the invention the server includes two stages of writing prior to committing to a write to the physical medium. This speeds both read and write access to the disk, allows concurrent user access to the disk, and greatly simplifies the mapping and file structures on the disk. In an alternate  
5 embodiment of the invention the server includes a hard drive for configurable caching of selected ones of the optical storage media.

FIG. 1 shows a plurality of devices coupled across the network with an optical storage server. Computers for 118, 120 and scanners 124,134 interface across the network 100 with the optical storage server 102. The optical storage server includes a  
10 server portion 106 and one or more optical storage drives, e.g. CD drives 108,110, and 112. Optical disks 104 may be inserted and removed from any or all of the above-mentioned drives.

There are three phases to the operation of the server: a configuration phase, a boot phase, and an operational phase. During the configuration phase a selected user  
15 with the appropriate administrative pass code can access html pages provided by the server to configure each of the drives in the server. In the example shown an administrator may access the server via a browser on computer 120 for example. After inputting of the appropriate network password the user is provided with HTML pages for configuring the individual drives on the server. The configuration options include:  
20 directory caching, directory and data caching, and archiving. Directory caching provides the ability to browse all discs in a changer without having to cycle through them to read the directories. Not until a file is accessed is there be a need for the changer to load the disc into the drive reader. Directory and data caching is the ability for CD-ROM disks to be mirrored (copied) to a hard disk co-located in a CD-ROM  
25 tower with CD or DVD drives. The data that has been copied is now accessible directly over the network from the hard disk, rather than from the CD itself. The primary advantage of this is access speed. When the user clicks on a file icon, the response back to the user is faster than if the file was being pulled from the CD directly. This is very apparent in the case where the CD drive has stopped spinning or  
30 "idled down" because it has not been used for a period of time. Data access from a CD requires the drive to spin up and then find the file which could take several seconds.

[illegible]

**Figure 1**

[illegible]



5

10

20

AXISP001

In UDF the primary data structure which defines and filters how a disk image can use files is called the file entry information control block (ICB). In UDF every file and directory has its own ICB. File data is generally written before its defining ICB is written, so if a file is located on one more than one extent, the ICB can contain a list of those extents. An ICB may point to its "parent" or an older version of the ICB. The second of these control structures is called a virtual allocation table (VAT) which is used to assign a sequential number of virtual reference to each file. Rather than referencing the logical address, the actual location, of each file or directory on disk as an ISO 9660, UDF can reference a virtual address. This method of referencing files is called indirection. These VAT can be placed on any location in the track but are referenced by a VAT ICB which is always placed on LWA on the media so it's easy for the CD recorder to locate. Even if a file is changed in some way, the entire series of file pointers need not change; only the VAT has to change. The VAT itself may be located on multiple extents. The ICB is small enough to exist on a single block and contains a list of these extents. Where a single data file is changed, deleted or rewritten it is not necessary to rewrite the entire VAT, rather just that portion of the VAT that references the changed file needs to be rewritten along with the VAT ICB. In the example shown the VAT and the VATICB will be expressed as a single data structure for simplification.

ICB 304 and 306 are generated by processor 206 utilizing above discussed program code. ICB 304 tracks the packets received from scanner 134, identifying them as extents. Within the corresponding ICB each sequential extent is assigned: a sequential number; a starting address identifying the location of the ICB in either of the stages or on the media; and a number corresponding with the size of the extent. Within ICB 304 the extent #1, with an address "1002", and a size of 50 bytes corresponds to the first packet from scanner 136 which is stored in the first stage along with the associated extent which contains the data to be written to the media. ICB 306 is associated with the packets from scanner 134. The first packet is accorded extent #1, the next available address in the first stage "1053", and a size of 40. The acceptance of these new packets corresponding to a new file or data stream to be written to the media results in the generation of a corresponding entry in the VAT 308 which is stored at this point in the volatile portion of the memory 208. Each entry in

09733707-120800

5

10

20

25

The data and control structures previously present in the 2<sup>nd</sup> stage are flushed to the media prior to the entry of the above mentioned data and control structures.

As is evident from all of FIGS. 3A-F and in particular the above discussed FIG. 3D the input of data, i.e. write requests, are handled serially. Data is initially written to the 1<sup>st</sup> stage. When that is full it is moved to the 2<sup>nd</sup> stage. Finally, when the second stage is full it is flushed, i.e. written to the storage media, e.g. CD/DVD. Data is not written over the network directly to either the 2<sup>nd</sup> stage or the media. By way of contrast data output, i.e. read requests may be handled directly from any one of the 1<sup>st</sup> stage, the 2<sup>nd</sup> stage or the storage media. Where a hard drive cache of data is implemented that data may be read directly from the hard drive as well. The processor 206 implements the logic determined by the program code stored in memory 208. This logic effects a the serial movement of data input to the server through the network module 200 through each of the stages to the storage media. The logic implemented by the processor also supports data output from the server through the network module of data to be read from whichever of the stages that data is stored in.

FIG. 3E shows the next phase of the write operation. The extents 138, 128, the ICB 306 and the corresponding VAT 308B have been flushed from the 2<sup>nd</sup> stage to the optical media over the address range 348-452. The LWA has been updated to the address “453”. The last of the extents/packets 140 from scanner 134 has moved from the 1<sup>st</sup> to the 2<sup>nd</sup> stage. Since this is the last of the packets associated with the DataStream from scanner 134 the corresponding ICB 304 is also written to the 2<sup>nd</sup> stage, followed by the VAT 308C.

FIG. 3F shows the completion of the write sequence with the flushing of the contents of the 2<sup>nd</sup> stage to the media. This could take place after an appropriate delay interval. The FWA is updated to address “655”. The associated processes for

the management of the 1<sup>st</sup> and 2<sup>nd</sup> stages and the optical media will be set forth in the following FIGS. 4AB.

The sequential staged write I/O method and apparatus discussed above in FIGS. 3A-F may be applied to write once or write many media in any of a number of formats without departing from the scope of the invention. It may be applied with and without a hard drive cache. A number of advantages are exhibited by the above discussed I/O apparatus and method: first, multiple users may concurrently read to or write from the storage medium; second, the aggregation and coalescing of data in the first and second stages saves valuable space on the CD. Less of the space associated with lead in and lead out sequences for each block written to the CD is required since block writes are delayed until the 2<sup>nd</sup> stage is full. Third two data extents belonging to the same file that are sequential can be coalesced in the first stage. This reduces the number of extents for which data structures need to be generated. For communications over the internet the latter feature is of significance due the fact that the typical network packet has payload/data portion limited to 1k. Where a sizeable file, an image file of several meg is being transferred, the ability to coalesce network packets into large extents before writing to the media greatly reduces both time spent writing to and reading from the media and also reduces the complexity of the data structures required to locate the file on the CD/DVD. Fourth, access time to the CD is greatly reduced because fewer write operations to the CD are required.

FIGS. 4A-B are process flow diagrams corresponding with the writing and reading of data to an optical or DVD media as shown in FIGS. 3A-F. Processing begins at start block 400 for which control is passed to decision process 402. In decision process 402 a determination is made as to whether a read request has been received. If so control passes to process 404. In process 404 the location of the data is determined using the VAT and FWA. Control is then passed to process 406. In process 406 the data is read from the location determined in the prior step. Control then returns to decision process 402. Where the next I/O request is not a read request control is passed to decision process 408. In decision process 408 a determination is made as to whether a timeout interval since the last write request has expired. If it has then control passes to process 410 for flushing of the data in the 1<sup>st</sup> stage to the 2<sup>nd</sup>

stage and the subsequent flushing of the 2<sup>nd</sup> stage in processes 446-450. These latter processes will be discussed in greater detail the following text. If a write request is received control is passed to decision process 412. In decision process 412 a determination is made as to whether the write request involves adding new data or the editing of existing data. If editing is involved control is passed to decision process 424. In decision process 424 a determination is made as to the location of the data for which editing is requested. If the data is in the 1<sup>st</sup> stage then control passes to process 426 in which the data is updated, added, or deleted. Next in process 428 an update of the corresponding ICB and other data structures is made to reflect the changes followed by a return to process 402 for the processing of the next I/O request. If alternately, in decision process 424 a determination is made that the requested data is located in the 2<sup>nd</sup> stage control passes to decision process 430. In decision process 430 a determination is made as to whether the request is for the replacement or addition of data. If the request is for replacement control passes to decision process 432 in which the data is replaced followed by a return to decision process 402 for the processing of the next I/O request. Alternately, if a determination is made that the request is for adding data to the 2<sup>nd</sup> stage an extent corresponding to that data is passed to processes 418 for the addition of the data to the 1<sup>st</sup> stage. Additionally, in process 434 the associated VAT and ICB entries are updated to reflect the addition. Where the optical media accepts multiple writes to the same address additional processes (not shown) may be added to allow the replacement of the data on the optical media. Alternately, the associated file structures, e.g. ICB and VAT may be updated to reflect the deletion of that data simply by alteration of the pointers on the media to skip the stale data without actual erasure of the data.

Returning to decision process 412, and a decision that the write request is for new data to be added to the media control is passed to process 418. In process 418 an appropriate ICB is created/updated to reflect the writing of the new extent/packet to the 1<sup>st</sup> stage. Next in process 420 the associated data within the packet is written as a corresponding extent to the 1<sup>st</sup> stage. Next, in decision process 422 a determination is made as to whether the 1<sup>st</sup> stage is full. If it is not control returns to decision process 402. If it is control passes to process 436. In process 436 using the least recently used (LRU) or other such algorithm data is flushed from the 1<sup>st</sup> stage to the 2<sup>nd</sup> and

corresponding updates are made to the associated ICB. Control is then passed to decision process 440. In decision process 440 a determination is made as to whether any of the flushed extent(s) are the last within the corresponding file/data stream. If so control passes to process 438 in which the corresponding ICB is moved to the 2<sup>nd</sup> stage as well. Then in process 442 corresponding updates are made to the associated data structures, e.g. ICB and VAT. Control is then passed to decision process 444 which is the same process executed directly after a determination in decision process 440 that the extent being flushed is not the last in the associated file/data stream.

Next in decision process 444 a determination is made as to whether the 2<sup>nd</sup> stage is full. If not control returns to process 402. If so control passes to process 446. In process 446 the up to date RAM resident copy of the VAT is copied to the 2<sup>nd</sup> stage. Then in process 448 the 2<sup>nd</sup> stage is flushed to the media and in process 450 the FWA is updated. Subsequently control returns to decision process 402 for the processing of the next I/O request.

FIG. 4B is a detailed process flow diagram of the processes associated with process block 404-406 discussed above. Processing begins at process block 460 in which the address to be read it is determined. Next in decision process 462 a determination is made as to whether the address to be read is greater than the LWA. If it is control passes to process 464 in which data is read from the 1st stage. Alternately, if the address is not greater than the LWA then in decision process 466 a determination is made as to whether it is less than the FWA. If it is then the requested data resides in and is read from the 2<sup>nd</sup> stage in process 468. If the results of decision processes 462, 466 are both negative then the data resides on the media 300 (See FIG. 3) and that data is located using the existing file pointers, e.g. VAT and ICB and is provided to the user.

Although the above discussed process flow are in some details specific to the UDF media format the two stage write architecture set forth in the above discussed embodiment may be applied with equal advantage to other media formats without departing from the scope of the invention.

FIGS. 5-7 show various aspects of a user configurable caching implementations that may be practiced on one or more device drivers coupled to the server 106. FIG. 5 is a hardware block diagram of the server, optical and electronic storage mediums shown in FIG. 1. The server 106 includes: a CD driver 212, a hard drive driver 210, the processor 206 and memory 208. Interfacing with the server through the associated drivers is a representative optical storage media device driver 108 with a CD 300 and a hard drive 114 with hard drive media 520. In the example shown the hard drive is formatted with the UDF file structure and includes requisite media identifiers and other directory structures 522. The FWA of the hard drive is "257" and the LWF is address "1000". The CD 300 is formatted with an ISO 9660 data architecture and includes a disk identifier 302 and a FWA of 257 and a LWA of 1000. Other formats may alternately be applied.

The information on the CD includes a plurality of extents 500, 502, 508, 512, and 516. The first of these extents 500 includes the associated ISO 9660 data structures including volume descriptors, directory structures, and the path tables. The second extent 502 includes two data files 504 and 506. The second extent 508 includes a single data file 510. The fourth extent 512 includes the data file 514. The last extent 516 includes a single data file 518.

In the example shown the user has selected directory and data caching for CD 300 on the hard drive media 520. This option is recorded in memory 208 after the user has selected via the above discussed browser interfaces the appropriate cache option. Directory and data caching involves the transfer 550 of the file structure and extent structure to the CD driver. These structures are revised and a directory for the CD is opened on the hard drive. Two files, an extent list 540 and a data file 542 are placed 552 in that directory as a result of that selection. If the user has selected directory caching only then only the extent list and file directories would be copied to the hard drive. The extent list 540 contains offsets for the extents on the hard drive, the data file 542 contains the file structure 530 and the individual files 504, 506, 510, 514 and 518. The space required for the files is considerably reduced due to the removal during the transfer of empty spaces or lead in/out segments which may be as



FIGS. 6A-C show the data structures associated with both the CD and hard drives shown in FIG. 5. FIG. 6A shows the CD 300 with the extents 500, 504 510 514 and 518. Empty space and or lead in and lead out segments separate several of the segments. The mapping of the CD extents E1-5 to extents H1-H4 on the hard drive data file 542 is shown. On the hard drive empty space is removed. Additionally the five segregated extents containing on the CD and the associated files stored in various portions thereof are coalesced to a single file 542 on the hard drive with four offsets. The coalescing of extents speeds up hard drive access, while the removal of empty and or lead-in/out space reduces hard drive size requirements. The use of a single file with offsets greatly speeds access to the hard drive as opposed to multiple files being written thereto.

AXISP001

either the configuration or operational phase of the server. These parameters may be changed at any time to update the cache policy for the associated CD/CD Drive. The selection of cache policy may be implemented using the browser based interfaces shown in FIG. 2.

5           FIG. 6C shows a directory tree on the hard drive 520. A dedicated directory 670 for CD 300 is shown. When directory and data caching are selected for the drive two files 540 and 542 for the extent list and data respectively will be created. When partial caching is implemented both files will be present on the hard drive but the data file 542 will only contain the meta data or file structures 530 (See FIG. 5). This will  
10       allow the user to browse the directory and file names for the associated CD from the hard drive. Access to the actual data requires an access to the CD.

          FIGS. 7AB are process flow diagrams corresponding with the boot phase and operational phase of operation of the server shown in FIGS. 5-6AC. Processing and begins at of the start block 700 from which control is passed to decision process 702.  
15       In decision process 702 a determination is made as to whether the next CD has been configured for any one of the various cash configurations discussed above. These include: directory caching, directory and data caching and archiving. This decision is made by the server on the basis of cache selections made by the user during system configuration and stored in memory 208 (See FIG. 2 ). If no cache option has been  
20       selected control passes to process 704 for the initialization of the next CD on the SCSI or other bus connected to the server. Alternately, if caching has been enabled control passes to process 706. In process 706 the directory for the selected CD is created on the hard drive. Then in process 708 the list of CD file extents is obtained from the CD data structures. Where the CD is formatted in an ISO 9660 format this information is  
25       obtained from the path table. Variations in these processes will apply where alternate formats for the CD are encountered. Next in process 710 additional extents are obtained from the volume recognition descriptors (VRD) on the CD. Then in process 712 the extents which are determined to exist on the CD-ROM and which are contiguous thereon are mapped into a single extent. This step reduces the amount of  
30       time required to subsequently access either the CD or the hard drive. In process 714 the extent list 540 (See FIGS. 5, 6B-C) is stored in the directory 670 for the CD 300

on the hard drive (See FIG. 6C). Next, in decision process 716 a CD data file 542 (See FIG. 6C) is created in the above discussed CD directory on the hard drive. The initial offsets for the hard drive extent list fields 662 (See FIG. 6B) are recorded in the extent list. The meta data or file structures 530 (See FIG. 5) for the CD are then  
5 copied from the CD into this file. The actual data is not copied at this time.

Next in decision process 718 a determination is made as to whether both directory and data caching has been enabled. If only partial caching is enabled control returns to process 704 for the processing of the next CD drive. Alternately, if directory and data caching has been enabled control passes to decision process 722.  
10 The filing processes 722 and 728 result in the calculation of all offset fields 662 by which a file on the CD may be mapped to the hard drive.

In decision process 722 the next extent is processed beginning with process 724. In process 724 data in the extent on the CD is copied to the hard drive at the prior recorded offset. Then control is passed to process 726 in which the next offset  
15 for the data in the hard drive is calculated using the prior offset value and the length of the data copied to the hard drive the preceding process. Control is then passed to process 728 in which the calculated offset is registered in the corresponding one of fields 662 in the extent list 540 (See FIG. 6B). Control then returns to decision process 722 for a determination of the existence of the next extent.

FIG. 7B shows the processes associated with read access to the particular CD. Processing begins at start block 750 and proceeds to decision process 752. In decision process 752 a determination is made as to whether a request for the display of a directory of a particular CD has been received by the server. When such a request is received control passes to decision process 754. In decision process 764 a  
20 determination is made as to whether directory caching has been enabled in which event control is passed to process 758. In process 758 the extent list and associated directory structures stored on the hard drive as a during partial caching are utilized to rapidly display to the user all of the directories and files structures for the CD. Control subsequently passes to decision process 760. Alternately in decision process 754 when  
25 a determination is made that partial caching has not been enabled than control is passed to process 756. In process 756 the disk for the selected CD is spun up and the extent

5

10

25

The many features and advantages of the present invention are apparent from the written description, and thus, it is intended by the appended claims to cover all such features and advantages of the invention. Further, since numerous modifications and changes will readily occur to those skilled in the art, it is not desired to limit the  
5 invention to the exact construction and operation as illustrated and described. Hence, all suitable modifications and equivalents may be resorted to as falling within the scope of the invention.

09/33/07 12:08:00